Center for Advanced Computation

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
URBANA, ILLINOIS 61801

CAC DOCUMENT NO. 218

ERROR ANALYSIS OF THE COMBINED ILLINOIS-BROOKHAVEN ENERGY MODEL

by

Clark W. Bullard, III

November 1976

CAC DOCUMENT NO. 218

ERROR ANALYSIS OF THE
COMBINED ILLINOIS-BROOKHAVEN ENERGY MODEL

by

Clark W. Bullard III

Energy Research Group
Center for Advanced Comput
University of Illinois at Urban
Urbana, Illinois 61801

November 1976

This work was supported by the Energy Researcl
Division of International Affairs

CAC DOCUMENT NO. 218

ERROR ANALYSIS OF THE
COMBINED ILLINOIS-BROOKHAVEN EN

by

Clark W. Bullard III

Energy Research Group
Center for Advanced Computation
University of Illinois at Urbana-Champaign
Urbana, Illinois 61801

November 1976

This work was supported by the Energy Research and Development Administration, Division of International Affairs

ABSTRACT

The effects of parametric uncertainty in the combined input-output/linear programming model are examined. The mathematical structure of the model is such that small errors on input parameters may be magnified through the solution process, resulting in errors of several hundred percent in some model outputs. Error bounds due to parametric uncertainty in the input-output submodel are evaluated, in terms of their impact on the combined model solution. Though significant, these errors are shown **much smaller than** those caused by uncertainty in the parameters of the linear programming submodel. As a result, extreme caution is called for in presenting and interpreting results of combined model calculations.

ACKNOWLEDGEMENT

This work was supported by the Energy Research and Development Administration, Division of International Affairs. Thanks are due to Lee Abramson for his support and interest in analyses of the effects of error magnification through mathematical models, and to John McMillan, for performing all computer programing for this analysis.

1.0 INTRODUCTION

The benefits obtainable from sophisticated models of energy flow through the economic system are well known. But such models are worthless, unless the user is confident that uncertainties on model results are within tolerable limits.

Errors are always present in observations of system characteristics, and are reflected as uncertainty in parameters of mathematical models of such systems. It is possible for model results to be extremely sensitive to small changes in certain sets of parameters. Depending on the mathematical structure of the model, input errors may be magnified more than a hundred-fold. It is therefore severity of this potential problem depends on the mathematical structure, so all necessary to subject such models to rigorous analysis before results are accepted.

One of the most detailed energy-economic models used by the U.S. Energy Research and Development Administration is a combined linear programming/input-output model developed by Brookhaven National Laboratory and the University of Illinois.* The model is used primarily for scenario analysis of alternative designs for energy supply and delivery systems. It is quite detailed, with about 110 energy supply and consumption technologies represented by over 10,000 parameters. The purpose of this report is to quantify the effects of parametric uncertainty on model results.**

1.1 Importance of the Problem

This problem is important for two reasons. The first is analytical.

---

* These models are described by Behling, et al. (1975) and Bullard and Sebald (1975).

** Other sources of errors affecting model results (e.g. linearity assumptions) are outside the scope of this report.

The two submodels were combined in order to reduce uncertainty in each; the linear programming (LP) submodel updates the fuel mix representation for the input-output (IO) submodel while the IO calculates a consistent set of input data for the LP, corresponding to given economic conditions. This paper presents a calculation of the uncertainty remaining in the combined model results after these two sources have been removed. The problem is also important for a professional reason, the analyst's responsibility to define and report error bounds. Through repeated application and hearsay, models are sometimes assigned much more credibility than the model-builder's caveats will justify. It seems inevitable that somewhere along the line from analyst - reviewer - supervisor - editor - advocate - policymaker, that some of the modeler's caveats will be separated from the table of results and relegated to "reference 1." The work reported here is an attempt, in the face of this inevitability, to analyze the sensitivity of a model to parametric uncertainty, apart from any specific application.

1.2 Nature of the Models

Statistics generated by government agencies form the basis for most of the model's parameters. Parametric uncertainty can arise from a number of sources, ranging from data reporting and processing errors to outright lying on census and survey forms. Such errors have been estimated for the parameters of the energy input-output (IO) model, and their effect on model results was reported by Bullard, et al. (1976). Results of that analysis are used here to evaluate the combined model.

Sensitivity analyses of the linear programming model have been more limited, primarily for methodological reasons. In a linear programming

model, it is almost trivial to compute the sensitivity of the objective function (in this case the total cost of energy production) to small changes in constraints. Changes in other parameters (e.g. capital and transport costs, energy conversion efficiencies) can sometimes be analyzed one at a time, but simultaneous variation is much more difficult.

Until very recently, the effects of simultaneous variation of parameters were not well known for either the LP or IO models. The problem remains unsolved for the Brookhaven LP model. This report is limited to analysis of the effects of IO model uncertainty on the combined IO-LP model results. The objective is to incorporate these uncertainties into the calculation in a way that permits combined model results to be accompanied by upper and lower bounds.

Results of sensitivity analyses on the IO model are reviewed in Section 2. The method for incorporating these results in the combined model is discussed in section 3, where a brief description of the connection between the IO and LP submodels is given. Results and conclusions are presented in section 4.

2.0 BACKGROUND

There are many techniques for estimating upper bounds on the effects of parametric uncertainty. The methods easiest to apply are much too conservative* and therefore yield little useful information. In recent years,

---

* In this context, a "conservative bound" means a loose bound -- an overestimate of an upper bound.

successive attempts have been made to bound the uncertainty of results of the IO submodel; each attempt tightened the bounds. Since the analysis of uncertainty in the combined model leans heavily on these earlier results, a brief summary is presented here.

2.1 The IO Submodel

The IO submodel is simply a system of N(~100 for this application) equations and unknowns.

$$\underline{X} = (\underline{\underline{I}}-\underline{\underline{A}})^{-1}\underline{Y} \qquad (1)$$

The dependent variable is the vector $\underline{X}$ of total outputs from each sector; the independent variable is the vector $\underline{Y}$ of final goods and services making up the GNP; and $\underline{\underline{A}}$ is a matrix of parameters, A_{ij}, a typical element of which denotes the direct inputs of product i needed to produce a unit of product j. These parameters have been derived from economic statistics by the federal Bureau of Economic Analysis* for the base year 1967.

2.2 Maximum Error Bounds

The simplest method for estimating upper bounds on elements of $(\underline{\underline{I}}-\underline{\underline{A}})^{-1}$ due to uncertainty in $\underline{\underline{A}}$ is to calculate the condition number of $(\underline{\underline{I}}-\underline{\underline{A}})$ from its eigenvalues. This bound was found by Bullard and Sebald (1975) to be excessively conservative. In that paper, a tighter set of bounds was calculated, based on the assumption that values of all parameters $\underline{\underline{A}}$ were at their maximum error bounds and combined in such a way as to maximize errors in $(I-A)^{-1}$. Unfortunately, these results were also too conservative. For example it was found that a 10% uncertainty in elements of A produced

* See U.S. Department of Commerce, Bureau of Economic Analysis (1974).

maximum errors in $(I-A)^{-1}$ averaging in the 40-60% range. Obviously, one does not need a 10,000 parameter model to estimate next year's energy demand within ±40-60%.

2.3 Stochastic Sensitivity Analysis

The likelihood of all 10,000 parameters in $\underline{\underline{A}}$ being at their maximum upper bounds simultaneously is clearly small. It was expected on intuitive grounds that there would be a significant amount of internal error cancellation, as random errors offset one another. No analytic solutions to sensitivity analysis problem are available, and numerical solutions have not been possible until recent years due to computer hardware and software limitations. A numerical analysis has now been performed on this model, however, and the results show a considerable tightening of error bounds.* For example, the (three standard deviation) bounds on most elements of X are less than 10% , based on actual estimates of base year parametric uncertainty by the Bureau of Economic Analysis. It will be shown below that it is the uncertainty in energy sector elements of X that is critical to the functioning of the combined model.

3.0 METHOD

The combined model interconnections are shown in fig. 1. The IO model takes the given final demand vector $\underline{Y}$ and calculates the corresponding vector $\underline{X}$ of total sector outputs. Eight elements of $\underline{X}$ represent total requirements for energy services: space heat, air conditioning, industrial process heat, etc. (See Table 1). These results become constraints for the

---

* Bullard, et al (1976)

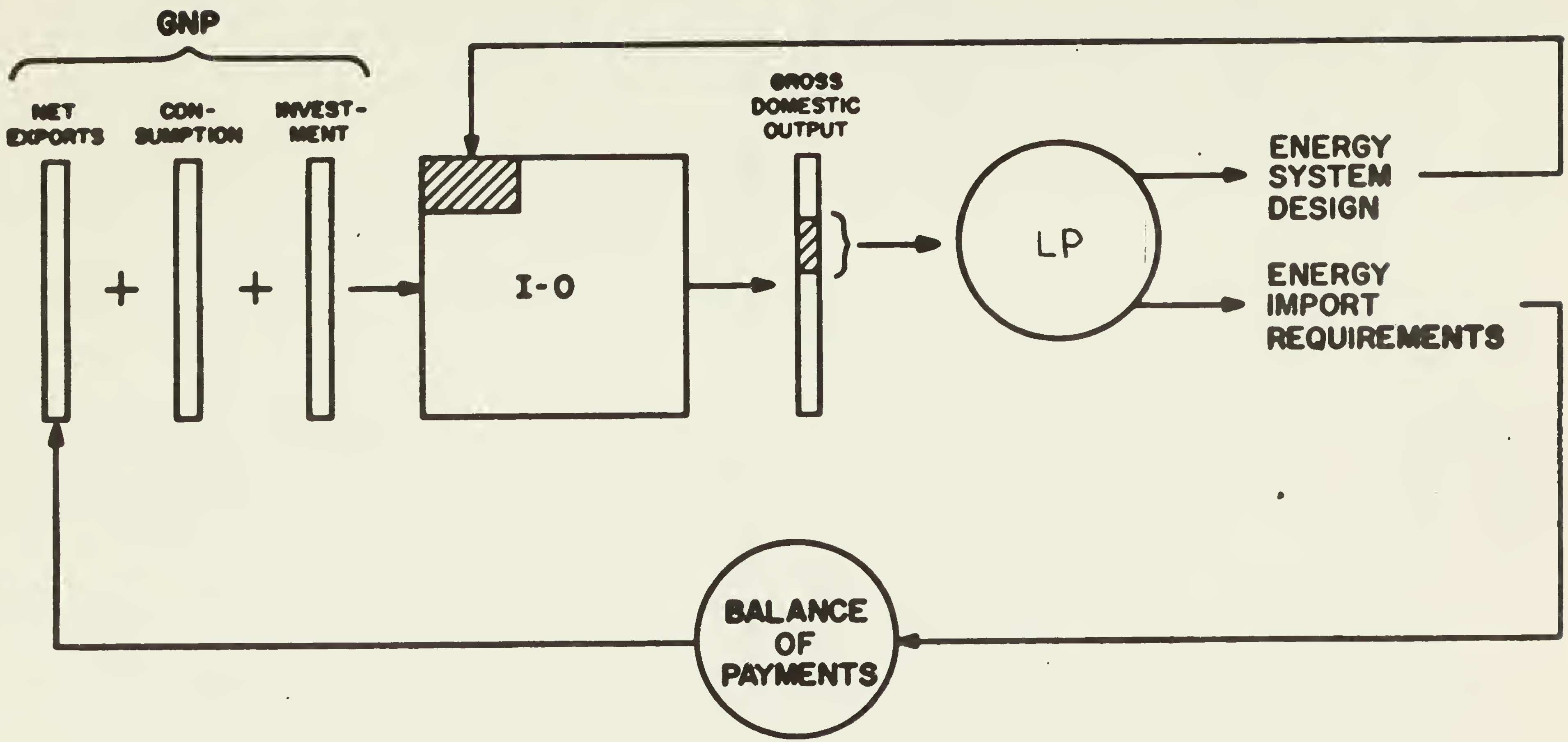

Figure 1. The Combined Input-Output/Linear Programming Model

Energy Service	Case I "1967" Uncertainty	Case II "1985" Uncertainty
Coke	2	4
Petrochemical feedstock	1	2
Motive power	2	4
Process heat	1	2
Water heat	7	14
Space heat	5	10
Air Conditioning	3	6
Electric Power	2	4

Table 1. Error Tolerances (% of Mean) for Total Energy Service Requirements

LP submodel, which then calculates the mix of energy supply and conversion facilities that minimizes its objective function (in this case total energy costs). These changes in energy supply technology (elements of a submatrix of $\underline{\underline{A}}$) are fed back to the IO model for another iteration. Convergence is usually achieved after 5 or 6 steps.

Due to uncertainty in $\underline{\underline{A}}$, all eight elements of $\underline{X}$ needed by the LP submodel will vary with each iteration. Therefore the shadow prices associated with each LP solution are not useful for estimating the response of the LP because of the simultaneous variation. Shadow prices only yield sensitivity information when only one constraint is varied within certain limits.

The maximum (3-sigma) error magnification factors obtained by Bullard, et al. (1976) were used to inflate the values of $\underline{X}$ after each iteration before they were passed to the LP submodel to be used as constraints. As bounds were calculated in this manner, convergence characteristics remained unchanged.* At each step, all 8 elements of $\underline{X}$ were set to their upper bounds; clearly not the most probable configuration. It was done to drive the LP in a somewhat consistent "worst case" direction, providing more energy resources to meet increased demand for each energy service. The combined model is such that an increase in any energy service demand brings about an increase in the objective function. However, these results will not necessarily bound the outputs of certain types of energy supply

---

*Actually, fig. 1 is slightly oversimplified; it does not show another minor connection between the IO and LP submodels. Certain technical coefficients in $\underline{\underline{A}}$ are used by the LP to define the transportation modal mix. These coefficients were set at their upper bounds also. The quantitative effect turned out to be small, so the slight double-counting (errors due to $\underline{\underline{A}}$ were already embodied in $\underline{X}$ uncertainty) was neglected.

and conversion facilities required for a particular scenario.*

In summary then, the "nominal" solution was first obtained using standard input data. Two "upper bound" solutions were then obtained using the results of IO stochastic analyses: Cases I and II. The next step was to assess the impact of this uncertainty on elements of the solution corresponding to critical design parameters of the energy system.

4.0 RESULTS AND CONCLUSIONS

Several cases were examined. In Case I, error bounds on X (total demand for energy services) were taken directly from the estimates of Bullard, et al. (1976) for errors inherent in the data for the model's base year, 1967. In Case II, larger error bounds were used to reflect the added uncertainty in model parameters due to technological change in the economic system during the period 1967-1985.** The LP submodel contained parameters estimated for 1985 based on 1970 data, and a GNP (market basket of final goods and services) estimated for 1985.*** These parameters and independent variables held constant for the two cases.

The error bounds on the 8 elements of X for Cases I and II are shown in Table 1. On the surface, these levels of uncertainty may appear small

---

*A large scale stochastic sensitivity analysis could provide this information, but might be prohibitively expensive. Further investigation could conceivably yield ways of using stochastic programming techniques in the LP submodel, but that is beyond the scope of this analysis.

**It is shown by Bullard, et al. (1976) that this corresponds roughly to a case where the uncertainty on each element of A is doubled (e.g. ±5% becomes ±10%).

***These were the latest estimates comprising the current data base for the combined model at Brookhaven National Laboratory in April, 1976. Tessmer (1976).

to some observers.* However it is the possibility that they may be considerably magnified that is of interest here.

Among the results analyzed were changes in key energy system design parameters. The most notable upper bound variations off the base case included (+67%, +222%) in coal steam electric capacity for Cases I and II, respectively. Concurrently, gas turbine power generation changed by (+9%, -72%), and domestic oil production increased by (+10%, +15%). These results clearly indicate the possibilities for magnification of input uncertainties through the LP submodel. It is well known that LP models are prone to flip-flop, in this case sending energy flows along one path or another. This sensitivity is critical to the central purpose of the model, which is to calculate energy system design parameters such as those identified above. Awareness of these potential problems has led users of the Brookhaven LP model to conduct extensive sensitivity analyses around "nominal" results.** The problem of course has been that of knowing what are reasonable bounds for the uncertainty on key parameters and constraints.

The error bounds on the energy service parameters, averaging about (3%, 6%) for Cases I and II, would introduce about the same magnitude and type of error as equivalent uncertainty in the GNP estimate for 1985. If the nominal estimate had been that GNP would grow at 3% over the next 10 years, realized growth rates of (3.3%, 3.6%) would introduce approximately the same errors in the LP submodel as did the errors in energy service demands of Table 1. Said another way, the cumulative uncertainty on 10,000

---

* Remember, this is the uncertainty due only to specification of the technology of producing goods and services in 1985. The GNP and market basket of goods is assumed to be known exactly.

** See Cherniavsky (1974) and Hoffman (1974).

technological parameters in the IO submodel amounts to less than that due to a 3-6% uncertainty in the level of 1985 GNP.*

It was stated earlier that while the "improbable" distribution of maximum errors will probably have a maximum impact on the objective function, they will not necessarily produce the largest uncertainties energy system design parameters: the most important outputs of the combined model. Nevertheless, the errors are substantial, and demonstrate the potential for considerable magnification of IO uncertainties through the LP submodel. The GNP uncertainty would most likely cause a nearly uniform errors on the energy service constraints, while the IO parametric errors would be randomly positive or negative. The statistical combination of these uncertainties could probably be represented by a ±10% variation on each constraint. For 1985 applications, this is not likely to be worse than many other LP constraints. The sensitivity of the LP submodel to such variations in all its constraints might be analyzed separately using this information.

Intuitively, the most important parameters determining the optimal fuel mix are the relative cost figures in the objective function of the LP submodel. The uncertainty surrounding 1985 relative energy prices is perhaps the largest in the entire model, Certainly 1965 estimates of 1975 prices would have missed the mark by a margin wide enough to render useless almost any 1965 estimates of "optimal" energy system design for 1975. Between now and 1985 pollution control costs, developments in scrubber technology, clean air act amendments, and coal transport costs are highly uncertain,

---

* This statement must be qualified to the extent that the simultaneous specification of maximum error bounds on energy service requirements was assumed.

but decisions must be made soon on coal fired power plants to go on line in 1985. Similarly, the outcome of referenda and unforseen costs in the nuclear fuel cycle cloud estimates of future costs of that technology. This is not new; the purpose here is to compare the effect of uncertainties in those LP model parameters to the effect of parametric uncertainty in the IO model.

Consider first the case that the estimated 1985 cost of electricity from coal turned out to be 5% too high. Ceteris paribus, the combined model results yield an optimal energy system having a 1500% increase in the number of coal fired power plants over the 1985 base case! There would be no nuclear electricity produced under these conditions, according to the model. A simple 2% increase in the relative cost of nuclear power produced the same result. The obvious conclusion is that for practical purposes, coal and nuclear power are "perfect substitutes" under the assumptions built into Brookhaven's 1985 base case calculations for ERDA. The point is that the possible error due to uncertainty in the 10,000 parameters of the IO model (recall they were +67%, +222% on coal electric production) is small compared to the possible errors from uncertain predictions of the future price of nuclear electricity.

In conclusion it must be said that relatively little is known about the relative importance of parametric uncertainty on the results of large complex models being used to develop U.S. energy policy. The linkages inside and between the various models have the potential for magnifying uncertainties on input data to a startling degree. As for understanding the effects of simutaneous stochastic variation of input data, the state of the art is in its infancy. Only recently did advances in computer hardware and software development make it possible to solve the stochastic error

problem for the IO submodel. In the absence of a similar analysis of the LP submodel, I have attempted here to do only two things: 1) roughly indicate how uncertainty in IO parameters can be magnified through solutions of the IO and LP submodels, and 2) point out the necessity for performing similar analyses of simultaneous variation of LP parameters. Only after completion of such systematic analyses, can a set of simple, workable guidelines be established for the interpretation of combined model results.

5.0 REFERENCES

D. Behling, W. Marcuse, M. Swift, R. Tessmer, "A Two-Level Iterative Model for Estimating Interfuel Substitution Effects" Summer Computer Simulation Conference, San Francisco, 1975.

C. Bullard, D. Amado, T. Sebald, D. Putnam, "Stochastic Analysis of Uncertainty in a U.S. Input-Output Model, CAC Document No. 208, Center for Advanced Computation, University of Illinois at Urbana, September 1976.

C. Bullard, T. Sebald, "A Model for Analyzing Energy Impact of Technological Change" Summer Computer Simulation Conference, San Francisco, 1975.

E. Cherniavsky, "Brookhaven Energy System Optimization Model" BNL 19569, Brookhaven National Laboratory, December 1974.

K. Hoffman, "Unified Framework for Energy System Planning," Energy Policy, Special Issue on Modeling, 1974.

R. Tessmer, personal communication, April 1976.

U.S. Department of Commerce, Bureau of Economic Analysis, "Input-Output Structure of the U.S. Economy: 1967," v.1-3, U.S. Government Printing Office, 1974.